arXiv:physics/0608183v1 [physics.soc-ph] 17 Aug 2006

Superiority of the h-index over the Impact Factor for Physics

Casey W. Miller
Department of Physics, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093, USA

Focusing specifically on physics periodicals, I show that the journal Impact Factor is not correlated with Hirsch's h-index. This implies that the Impact Factor is not a good measure of research quality or influence because the h-index is a reflection of peer review, and thus a strong indicator of research quality. The impact gap between multidisciplinary journals and physics-only journals is significantly reduced when h is used instead of the Impact Factor. Additionally, the impact of journals specializing in review articles is inherently deflated using h because of the limited number of annual publications in such periodicals. Finally, a reordering of the top ranking journals occurs with h when only the physics articles of multidisciplinary journals are considered, falling more in line with the average physicist's interpretation of a journal's prestige.

The journal Impact Factor (IF) has earned the adjective *notorious* for several reasons[1,2,3]. First, in contradiction to healthy philosophies of science, the data used to determine the IF are not publicly available. Second, authors have reported unethical editorial practices that have the motive of increasing a journal's IF, including editors insisting that additional references to their journal be added to a bibliography *before* accepting an article for publication. Third, publishing in high-IF journals is used as a filter for hiring tenure-track faculty and advancement to tenure. R. Monastersky's "The number that's devouring science" highlights other intriguing issues[1], and astutely notes that the pragmatic motto "publish or perish" has mutated to "publish in a high-impact journal or perish."

The impact factor is, in spirit, a reasonable metric: the total number of citations divided by the total number of articles[4]. The three highest IF periodicals in which physicists publish were reported for 2003 as *Nature* (31.0), *Science* (29.2), and *Reviews of Modern Physics* (28.2)[5]. Many physicists find it odd that *Physical Review Letters*, historically the most well respected physics journal, boasts a meager 7.0. Another indication that the IF rankings are questionable is that experts disagree with the rankings for their subfields: few of the top titles are recognized, and even fewer are considered prestigious. Upon inspection, most of the top IF journals in each subfield specialize in review articles. Review articles are typically highly cited, though only a few articles are published annually. This causes such journals' high (and arguably, anomalous) IF, while simultaneously making original research journals appear relatively weak.

Based on the IF's several peculiarities, alternative measures of journal quality are being sought. In my opinion, the only reasonable way to rank physics journals is for physicists to measure the quality of published physics in each journal. While this may at first seem unrealistic, an indirect form of this type of peer review already exists in the form of citations. Articles of high quality or broad interest receive many more citations than articles of low quality or limited interest, and are therefore more scientifically influential. Extending this idea to periodicals, one can inspect the citation and publication history of a journal in a scientific and unbiased way to determine the average impact of its publications. I thus propose the h-index as a logical measure of a journal's influence on science.

The h-index, developed by J. E. Hirsch to quantify the scientific research output of an individual[6], has become popular because it is logically sound, simple to understand, and, most importantly, simple to calculate with easily obtained data. An individual's h-index is determined by searching a scientific database like Thompson's ISI Web of Knowledge[7] for all articles by an author, and ranking the output articles by the number of citations such that article 1 has the most citations; h is the rank of the lowest ranking article whose number of citations is bounded below by its rank. One can arbitrarily extend this procedure to journals, departments, institutions, or even zip codes. A graphical definition of the h-index is shown in the inset of Fig. 1. If $c(p)$ is the number of citations for paper p in the ordered list, then h is the intersection of $c(p)$ with the line $c'(p) = p$. I have empirically observed that $c(p)$ is bounded below by a right isosceles triangle with legs of length $2h$, such as ABC; violations are more likely to exist for scientifically young subjects due to inadequate statistics. An (approximately) equivalent definition of h is the coordinate of the intersection of the hypotenuse BC with $c'(p) = p$.

The total number of citations can be used to develop a related, but arguably more poignant index. Consider ADE, a similar triangle of ABC, whose area (α) equals $\sum_p c(p)$. Using the geometric definition, a new index, which I cannot resist naming the $\hbar$-index, is the coordinate of the intersection of the hypotenuse $\underline{DE}$ with the line $c'(p) = p$, or mathematically, $\hbar = \sqrt{(\alpha/2)}$. The $\hbar$-index is a more comprehensive measure of the overall structure of $c(p)$ for two reasons. First, $\hbar$ incorporates the most highly cited articles, while h basically ignores all articles with citations much greater than h. Second, $\hbar$ takes into account the body of articles with moderate numbers of citations, while h again ignores all such articles.

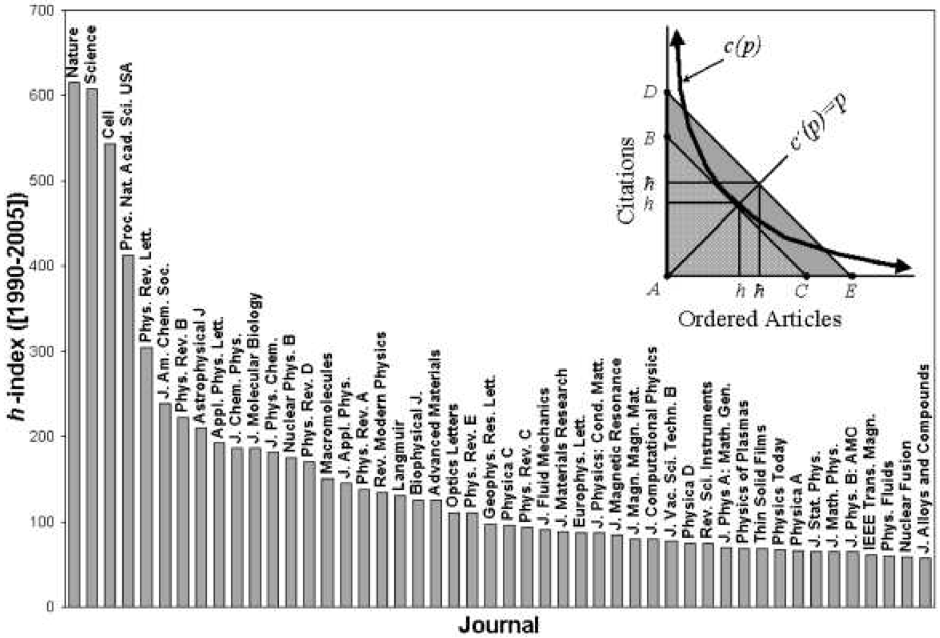

FIG. 1: (Color online) h-index for fifty journals from major subfields of physics. (Inset) Graphical representation of the indices h and $\hbar$ as defined in the text.

To illustrate the main difference between h and $\hbar$, consider two similarly aged[11], fully tenured, condensed matter physicists from the University of California, San Diego: L. J. Sham and I. K. Schuller. These professors are both well respected in this field, and have each received numerous awards for their scientific achievements. Both have $h = 54$, indicating that they have contributed approximately the same amount to science. However, h does not take into account that Sham's article with W. Kohn[12] that introduced density functional theory−for which Kohn was awarded the Nobel Prize in 1998−is the most highly cited paper in the history of *Physical Review*. Professor Schuller, on the other hand, has a large body of work and is among the most highly cited researchers in physics, though his single most highly cited paper has an order of magnitude fewer citations than Sham's (and neither he nor his immediate coworkers has yet to win the Nobel Prize). I find that Sham has $\hbar = 102$, and Schuller has $\hbar = 76$. Thus the *Sham-Schuller Paradox* is resolved using $\hbar$. Degeneracies with the $\hbar$-index will of course exist, but these will not be as qualitatively displeasing as in this and similar cases.

For some of the individuals noted in Hirsch's original article, I find the following for the indices $(h, \hbar)$: E. Witten (112, 163), P. W. Anderson (96, 164), S. Weinberg (89, 139), J. N. Bahcall (77, 102), D. J. Scalapino (76, 95), S. G. Louie (76, 97), R. Jackiw (69, 106), C. Vafa (67, 82), D. J. Gross (67, 106), and S. W. Hawking (62, 98).

Applying the h-index to individuals has proven very effective. Hirsch unequivocally showed that a large h, and more importantly $\mathrm{d}h/\mathrm{d}t$, indicates a successful scientist. Additionally, a recent hind-sight study showed a strong correlation between h and committee peer review: individuals that were granted prestigious post-doctoral fellowships[8] in biomedicine by a committee of well known scientists in that field from 1990 to 1995 had on average higher h-indices than other applicants[9]. Upon reflection, this study indicates the obvious: h measures how one's contributions are viewed by one's peers. Highly valued articles will receive many citations from the scientific community, which will result in a higher h-index for the authors or publishing journal. Similar correlation of h with peer review was recently demonstrated in chemistry[10].

I compared the rankings of scientific periodicals frequently targeted by physicists from a variety of subfields using h and $\hbar$ to evaluate the *IF* as a measure of research quality. The data used to calculate both indices were obtained using ISI to search by "source", limited to 1990-2006, and the following document types: Article, Letter, Review, Correction, Editorial Material, or Note. The latter three were included because comments, errata, and retractions are listed under these headings. Journals that began publishing after 1990 were excluded[13]. I chose sixteen years as an estimate of the time between entering graduate school and receiving tenure at a university. Additionally, this long timescale reduces the influence of high frequency fluctuations, such as those due to spectacular claims that are often

TABLE I: h, $\hbar$, 2003 Impact Factor, total number of publications, and the percent of uncited articles as of late March, 2006 for the subset of a journal's publications noted in the text. [a]Based on title. [b]Based on automated filter[16].

JOURNAL	h	$\hbar$	*2003 IF*	TOTAL	UNCITED
Nature	616	1157	31.0	39322	32
Science	608	1110	29.2	34361	30
Cell	543	864	26.6	7030	5
Proc. Natl. Acad. Sci. U.S.A.	413	1175	10.3	45192	6
Phys. Rev. Lett.	305	869	7.0	46461	8
Nature (only physics)	263[a], 285[b]	499[b]	-	-	-
Science (only physics)	260[a]	-	-	-	-
J. Am. Chem. Soc.	239	815	6.5	39513	6
Phys. Rev. B	223	796	3.0	77331	12
Astrophys. J.	210	654	6.6	35318	6
Appl. Phys. Lett.	193	615	4.0	44480	13
J. Chem. Phys.	187	623	3.0	39701	9
J. Mol. Biol.	187	-	5.2	12898	7
J. Phys. Chem.	181	646	3.3	47702	11
Nucl. Phys. B	175	386	5.3	13888	28
Phys. Rev. D	170	484	4.6	27902	10
Macromolecules	151	475	3.6	20954	7
J. Appl. Phys.	146	493	2.2	44925	15
Phys. Rev. A	138	426	2.6	26202	12
Rev. Mod. Phys.	134	188	28.2	573	6
Langmuir	130	-	3.1	18201	11
Biophys. J.	126	333	4.5	9572	9
Adv. Mater.	125	-	7.3	4659	15
Opt. Lett.	111	280	3.4	11073	13
Phys. Rev. E	111	363	2.2	26660	16
Geophys. Res. Lett.	97	-	2.4	16313	16
Physica C	96	-	1.2	13567	18
Phys. Rev. C	93	-	2.7	13407	11
J. Fluid Mech.	91	-	1.8	5958	11
J. Mater. Res.	88	-	1.6	7141	15
Europhys. Lett.	87	-	2.1	7872	14
J. Physics: Cond. Mat.	87	-	1.8	18579	18
J. Magn. Resonance	85	-	2.1	3374	12
J. Magn. Magn. Mat.	80	-	0.9	18138	30
J. Comput. Phys.	79	-	1.8	3835	16
J. Vac. Sci. Technol. B	77	-	1.6	10213	18
Physica D	75	-	1.6	2673	8
Rev. Sci. Instrum.	75	-	1.3	13959	23
J. Phys A: Math. Gen.	70	-	1.4	13144	20
Phys. Plasmas	69	-	2.1	7672	19
Thin Solid Films	69	-	1.6	17023	22
Phys. Today	67	-	5.0	5652	81
Physica A	66	-	1.2	5374	15
J. Stat. Phys.	65	-	1.2	3652	18
J. Math. Phys.	65	-	1.5	6975	23
J. Phys. B: AMO	65	-	1.7	8493	12
IEEE Trans. Magn.	61	-	1.0	16283	29
Phys. Fluids	60	-	1.6	4924	20
Nucl. Fusion	59	-	3.4	2914	15
J. Alloys and Compounds	57	-	1.1	12795	25
J. Physics D: Appl. Phys.	55	-	1.3	7597	20
Plasma Phys. Contr. Fusion	49	-	2.8	2946	18

disproved soon after their initial publication (the *IF* only uses two years[4]). The data were harvested in a 10 day window in late March, 2006; results are presented in Table I.

Figure 1 shows the h-index of fifty physics-related periodicals. As I discuss below, the relatively large h of the top four is due to the fact that these periodicals publish articles from other scientific disciplines, which inflates their indices relative to physics-only journals. *Physical Review Letters* also has a jump relative to the remaining physics journals. This is probably because it publishes work from all of physics, while the others are for specific subfields.

Figure 2 shows the journals ranked using $\hbar$, h, and the 2003 *IF* as published in the Journal Citation Report[5] for twenty titles with $h > 100$. The arrows show how the individual journals change rank when analyzed with the different indices. A journal's rank changes by an average of 1.3 positions between $\hbar$ and h, 4.4 between $\hbar$ and *IF*, and 3.4 between h and *IF*. The *IF*-based rankings are inconsistent with h and $\hbar$. *This means the IF is inconsistent with peer review, and is direct evidence that the IF is not a good measure of the quality of physics published in a particular journal.* Interestingly, the greatest leap of -17 places was made by *Reviews of Modern Physics*. This may indicate that

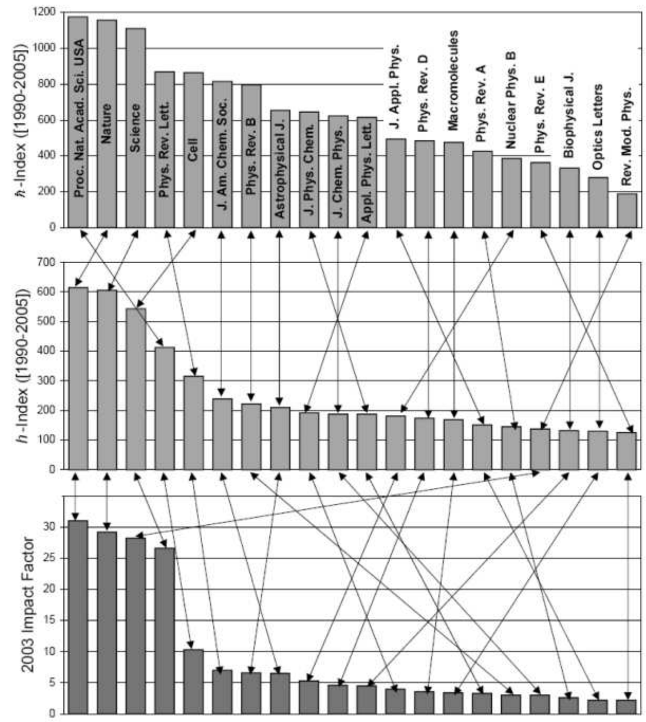

FIG. 2: (Color online) $\hbar$, h, and the 2003 *IF* for twenty physics journals with $h > 100$. The arrows show the re-ordering between the indices.

h intrinsically solves the impact inflation problem of review journals.

The large discrepancy between indices for *Nature* and *Science* relative to physics-only journals is due to the fact that their indices are inflated by work from other scientific fields[14]. Evidence of this inflation includes: 1) I can only identify $\sim 8\%$ of the top h articles in *Nature* and *Science* as physics research based on a qualitative assessment of their content by their titles[15]. 2) Physics articles represent only a small fraction of the total number of articles in these periodicals, which means fewer physics articles per unit time. The total number of articles is important in determining h and $\hbar$. Fewer articles means fewer citations, and thus an inherently lower h-index (as seen with *Reviews of Modern Physics*). Together, these points indicate that the impact of multidisciplinary journals, regardless of the metric, is heavily weighted by topics other than physics.

An important question naturally presents itself: *how does the subset of physics articles in multidisciplinary periodicals compare with physics-only journals?* To investigate this, an automated filter was used to select "physics" articles from the aforementioned data set of *Nature* publications. The selection criteria were based solely on the references of each article. An article initially qualified as physics if it contained at least one physics reference. The percentage of an article's total references that qualified as physics references served as a final, tunable filter. A reference was deemed a "physics

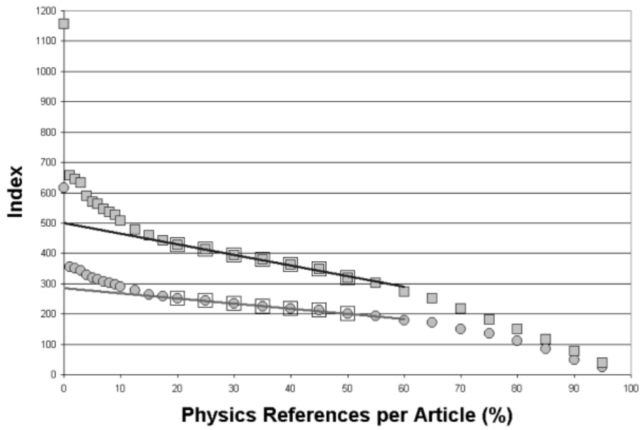

FIG. 3: (Color online) h (lower, circles) and $\hbar$ (upper, squares) indices for *Nature* using the percent of physics references in individual publications as a filter. The highlighted data were used for the linear fit, which yields $h_{phys} = 285$ and $\hbar_{phys} = 499$ when extrapolated to the ordinate-intercept.

reference" if it contained generic search strings (`astron`, `biophys`, *etc.*) or specific abbreviations (`phys fluids`, `j phys b`, *etc.*), and did not contain other strings (`physio`, `rehab`, *etc.*)[16]. The filter excluded `nature` and `science` because a disproportionate number of references contain at least one of these strings, which would have undermined this analysis. This method is certainly not perfect, but it is robust in the sense that perturbing the search strings does not significantly alter the results; the indices change by less than five percent when `phys rev` is purposefully excluded with the filter.

Figure 3 shows h and $\hbar$ for the qualifying *Nature* articles as a function of the final filter percentage. Each point on these curves was obtained by 1) selecting all articles whose bibliographies contained *at least* n% physics references, then 2) determining h and $\hbar$ for this subset of articles. The tail below fifteen percent is primarily due to physics references in non-physics articles. Both indices decay with increasing filter strictness as a result of the decreasing number of qualifying articles. The data around the inflection points (20-50%) were fit to lines, and the ordinate-intercepts were used to define physics-specific indices. I thus determined $\hbar_{phys} = 499$, and $h_{phys} = 285$. The latter is in good agreement with the manual analysis based on article titles: $h_{phys} = 263$ ($h_{phys} = 260$ for *Science*). With $\hbar = 869$ and $h = 305$, *Physical Review Letters* ranks higher than *Nature* (and presumably *Science*) for both indices. The origin of the large discrepancy between $\hbar_{phys}$ and $\hbar$ for *Nature* and *Physical Review Letters*, respectively, is probably due to the large percentage of uncited articles in *Nature* (see Table I).

In summary, the journal *IF* was shown to be inconsistent with the h-index. Insofar as h reflects peer review, and peer review reflects research quality, these results indicate that the *IF* is a poor measure of research quality. An additional benefit of h is its intrinsic grounding of the impact of review journals, whose *IF* is anomalously enhanced by their few annual publications. I showed that the impact of multidisciplinary journals is enhanced relative to physics-only journals by the more numerous non-physics articles contained therein. An analysis of the physics subset of *Nature* and *Science* revealed that *Physical Review Letters* has a greater h than either of these article subsets. This is in accord with *Physical Review Letters*' historical status as the most reputable physics journal. Based on these many observations, I conclude that the indices h and $\hbar$ are superior to the journal *IF* as indications of the quality of research published in a journal.

Acknowledgments

Special thanks to M. D. Chabot, T. Gredig, Z.-P. Li, T. C. Messina, and W. F. Egelhoff, Jr., for useful comments. This investigation was performed in the author's spare time, and was not supported by any funding agency.

---

[1] chronicle.com/weekly/v52/i08/08a01201.htm.
[2] chronicle.com/weekly/v52/i08/08a01701.htm.
[3] chronicle.com/colloquy/2005/10/impact/.
[4] chronicle.com/weekly/v52/i08/08a01201.htm#impact.
[5] www.research.hbi.ir/impact/Journal%20Citation%20 Report%20-%202003.htm.
[6] J. E. Hirsch, Proc. Nat. Acad. Sci. USA **102**, 16569 (2005).
[7] isiknowledge.com.
[8] One must of course cautiously apply h, or any such metric, to young researchers since it will be based on only a few papers, and is undoubtedly more a reflection of pedigree than potential.
[9] L. Bornmann and H.-D. Daniel, Scientometrics **65**, 391 (2005).
[10] A. F. J. van Raan, Scientometrics **67**, 491 (2006).
[11] This point is disputed by IKS, but it is accurate to $\sim 15\%$.
[12] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133 (1965).
[13] Exceptions: J. Phys. Chem. A and B, the fission products of J. Phys. Chem.; all three are included under the parent name of J. Phys. Chem., using a logical OR to search. J. Fluid Mech. started in 1991.
[14] *Cell* and *Proc. Nat. Acad. Sci. USA* are neglected here because their scope is more limited than *Nature* and *Science*; their high *IF*s are likely a reflection of the relative physics and life-sciences populations (~1:4, based on graduate degrees awarded per year in the USA, as per http://www.nsf.gov/statistics/wmpd/graddeg.htm).
[15] Interestingly, a few topics dominated the top h articles: nanotubes or fullerenes (34% for *Nature*, 32% for *Science*), superconductivity (16% in *Nature*), and magnetism (15% in *Science*)−the remaining physics articles were divided among other various topics.
[16] http://www.physics.ucsd.edu/~cmiller/filter.htm.